

Syntactic Complexity of \mathcal{R} - and \mathcal{J} -Trivial Regular Languages^{*}

Janusz Brzozowski, Baiyu Li

David R. Cheriton School of Computer Science, University of Waterloo
Waterloo, ON, Canada N2L 3G1
{brzozo, b5li}@uwaterloo.ca }

Abstract. The syntactic complexity of a regular language is the cardinality of its syntactic semigroup. The syntactic complexity of a subclass of the class of regular languages is the maximal syntactic complexity of languages in that class, taken as a function of the state complexity n of these languages. We study the syntactic complexity of \mathcal{R} - and \mathcal{J} -trivial regular languages, and prove that $n!$ and $\lfloor e(n-1)! \rfloor$ are tight upper bounds for these languages, respectively. We also prove that 2^{n-1} is the tight upper bound on the state complexity of reversal of \mathcal{J} -trivial regular languages.

Keywords: finite automaton, \mathcal{J} -trivial, monoid, regular language, reversal, \mathcal{R} -trivial, semigroup, syntactic complexity

1 Introduction

The *state complexity* of a regular language is the number of states in the minimal deterministic finite automaton (DFA) accepting that language. An equivalent notion is *quotient complexity*, which is the number of distinct quotients of the language. The *syntactic complexity* of a regular language is the cardinality of the syntactic semigroup of the language. Since the syntactic semigroup of a regular language is isomorphic to the semigroup of transformations performed by the minimal DFA of that language, it is natural to consider the relation between syntactic complexity and state complexity. By the *syntactic complexity of a subclass of regular languages*, we mean the maximal syntactic complexity of languages in that class, taken as a function of the state complexity of these languages.

Here we consider the classes of languages defined using the well-known Green equivalence relations on semigroups [15]. Let M be a monoid, that is, a semigroup with an identity, and let $s, t \in M$ be any two elements of M . The Green relations

^{*} This work was supported by the Natural Sciences and Engineering Research Council of Canada under grant No. OGP0000871 and a Postgraduate Scholarship.

on M are defined as follows:

$$\begin{aligned} s \mathcal{L} t &\Leftrightarrow Ms = Mt, \\ s \mathcal{R} t &\Leftrightarrow sM = tM, \\ s \mathcal{J} t &\Leftrightarrow MsM = MtM, \\ s \mathcal{H} t &\Leftrightarrow s \mathcal{L} t \text{ and } s \mathcal{R} t. \end{aligned}$$

If $\rho \in \{\mathcal{L}, \mathcal{R}, \mathcal{J}, \mathcal{H}\}$ is an equivalence relation on M , then M is ρ -trivial if and only if $(s, t) \in \rho$ implies $s = t$ for all $s, t \in M$. A language is ρ -trivial if and only if its syntactic monoid is ρ -trivial. In this paper we consider only regular ρ -trivial languages. \mathcal{H} -trivial regular languages are exactly the star-free languages [15], and \mathcal{L} -, \mathcal{R} -, and \mathcal{J} -trivial languages are all subclasses of star-free languages. The class of \mathcal{J} -trivial languages is the intersection of \mathcal{R} - and \mathcal{L} -trivial classes.

A language $L \subseteq \Sigma^*$ is *piecewise-testable* if it is a finite boolean combination of languages of the form $\Sigma^* a_1 \Sigma^* \cdots \Sigma^* a_l \Sigma^*$, where $a_i \in \Sigma$. Simon [18,19] proved in 1972 that a language is piecewise-testable if and only if it is \mathcal{J} -trivial. A *biautomaton* is a finite automaton which can read the input word alternatively from the left and from the right. In 2011 Klíma and Polák [10] showed that a language is piecewise-testable if and only if it is accepted by an acyclic biautomaton; here self-loops are allowed, that is, they are not considered cycles.

In 1979 Brzozowski and Fich [1] proved that a regular language is \mathcal{R} -trivial if and only if its minimal DFA is *partially ordered*, that is, it is acyclic as above. They also showed that \mathcal{R} -trivial regular languages are finite boolean combinations of languages $\Sigma_1^* a_1 \Sigma^* \cdots \Sigma_l^* a_l \Sigma^*$, where $a_i \in \Sigma$ and $\Sigma_i \subseteq \Sigma \setminus \{a_i\}$. Recently Jirásková and Masopust proved a tight upper bound on the state complexity of reverse of \mathcal{R} -trivial languages [9].

With regard to syntactic complexity, the following subclasses of regular languages were considered: In 1970 Maslov [12] noted that n^n was a tight upper bound on the number of transformations performed by a DFA of n states. In 2003–2004, Holzer and König [8], and independently Krawetz, Lawrence and Shallit [11] studied unary and binary languages. In 2010 Brzozowski and Ye [3] examined ideal and closed regular languages. In 2012 Brzozowski, Li and Ye studied prefix-, suffix-, bifix-, and factor-free regular languages [2]. In the same year, Brzozowski and Li [5] considered the class of star-free languages and three of its subclasses. Recently Brzozowski and Liu [6] studied finite/cofinite, definite, and reverse definite languages, where a language is *definite* (*reverse-definite*) if it can be decided whether a word w belongs to the language simply by examining the suffix (prefix) of w of some fixed length.

We state basic definitions and facts in Section 2. In Sections 3 and 4 we prove tight upper bounds on the syntactic complexities of \mathcal{R} - and \mathcal{J} -trivial regular languages, respectively. In Section 6 we prove the tight upper bound on the quotient complexity of reversal of \mathcal{J} -trivial regular languages, and we show that this bound can be met by our languages with maximal syntactic complexities. Section 6 concludes the paper.

2 Preliminaries

Let Q be a nonempty finite set with n elements, and assume without loss of generality that $Q = \{1, 2, \dots, n\}$. There is a linear order on Q , namely the natural order $<$ on integers. If X is a nonempty subset of Q , then the maximal element in X is denoted by $\max(X)$. A *partition* π of Q is a collection $\pi = \{X_1, X_2, \dots, X_m\}$ of nonempty subsets of Q such that

1. $Q = X_1 \cup X_2 \cup \dots \cup X_m$, and
2. $X_i \cap X_j = \emptyset$ for all $1 \leq i < j \leq m$.

We call each subset X_i a *block* in π . For any partition π of Q , let $\text{Max}(\pi) = \{\max(X) \mid X \in \pi\}$. The set of all partitions of Q is denoted by Π_Q . We can define a partial order \preceq on Π_Q such that, for any $\pi_1, \pi_2 \in \Pi_Q$, $\pi_1 \preceq \pi_2$ if and only if each block of π_1 is contained in some block of π_2 . We say π_1 *refines* π_2 if $\pi_1 \preceq \pi_2$. Then (Π_Q, \preceq) forms a poset. Furthermore, (Π_Q, \preceq) is a finite lattice; for any $\pi_1, \pi_2 \in \Pi_Q$, their *meet* $\pi_1 \wedge \pi_2$ is the \preceq -largest partition that refines both π_1 and π_2 , and their *join* $\pi_1 \vee \pi_2$ is the \preceq -smallest partition that is refined by both π_1 and π_2 . From now on, we simply refer to the lattice (Π_Q, \preceq) as Π_Q .

A *transformation* of a set Q is a mapping of Q into itself. In this paper we consider only transformations of finite sets Q . Let t be a transformation of Q . If $i \in Q$, then it is the *image* of i under t . If X is a subset of Q , then $Xt = \{it \mid i \in X\}$, and the *restriction* of t to X , denoted by $t|_X$, is a mapping from X to Xt such that $it|_X = it$ for all $i \in X$. The *composition* of two transformations t_1 and t_2 of Q is a transformation $t_1 \circ t_2$ such that $i(t_1 \circ t_2) = (it_1)t_2$ for all $i \in Q$. We usually drop the composition operator “ \circ ” and write $t_1 t_2$ for short. An arbitrary transformation can be written in the form

$$t = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ i_1 & i_2 & \cdots & i_{n-1} & i_n \end{pmatrix},$$

where $i_k = kt$, $1 \leq k \leq n$, and $i_k \in Q$. We also use the notation $t = [i_1, i_2, \dots, i_n]$ for the transformation t above. The *domain* $\text{dom}(t)$ of t is Q . The *range* $\text{rng}(t)$ of Q under t is the set $\text{rng}(t) = Qt$. The *rank* $\text{rank}(t)$ of t is the cardinality of $\text{rng}(t)$, i.e., $\text{rank}(t) = |\text{rng}(t)|$. The binary relation ω_t on $Q \times Q$ is defined as follows: For any $i, j \in Q$, $i \omega_t j$ if and only if $it^k = jt^l$ for some $k, l \geq 0$. Such a relation is indeed an equivalence relation, and each equivalence class is called an *orbit* of t . For any $i \in Q$, the orbit of t containing i is denoted by $\omega_t(i)$. The set of all orbits of t is denoted by $\Omega(t)$. Clearly, $\Omega(t)$ is a partition of Q .

A *permutation* of Q is a mapping of Q onto itself. In other words, a permutation π of Q is a transformation where $\text{rng}(\pi) = Q$. The *identity* transformation 1_Q maps each element to itself, that is, $j1_Q = j$ for $j = 1, \dots, n$. A transformation t is a *cycle* of length k , where $k \geq 2$, if there exist pairwise different elements i_1, \dots, i_k such that $i_1 t = i_2, i_2 t = i_3, \dots, i_{k-1} t = i_k$, and $i_k t = i_1$. A cycle is denoted by (i_1, i_2, \dots, i_k) . For $i < j$, a *transposition* is the cycle (i, j) . A *singular* transformation, denoted by $\begin{pmatrix} i \\ j \end{pmatrix}$, has $it = j$ and $ht = h$ for all $h \neq i$. A *constant* transformation, denoted by $\begin{pmatrix} Q \\ j \end{pmatrix}$, has $it = j$ for all i .

The set of all transformations of a set Q , denoted by \mathcal{T}_Q , is a finite semigroup, in fact, a monoid. We refer the reader to the book of Ganyushkin and Mazorchuk [7] for a detailed discussion of finite transformation semigroups.

For general definitions and facts about regular languages, we refer the reader to the handbook chapter by Yu [20]. Let Σ be a non-empty finite alphabet. Then Σ^* is the free monoid generated by Σ , and Σ^+ is the free semigroup generated by Σ . A *word* is any element of Σ^* , and the empty word is ε . The length of a word $w \in \Sigma^*$ is $|w|$. A *language* over Σ is any subset of Σ^* . The *reverse* of a word w is denoted by w^R . For a language L , its *reversal* is the language $L^R = \{w \mid w^R \in L\}$. The *left quotient*, or simply *quotient*, of a language L by a word w is the language $L_w = \{x \in \Sigma^* \mid wx \in L\}$.

The *Myhill congruence* [14] \approx_L of any language L is defined as follows:

$$x \approx_L y \text{ if and only if } uxv \in L \Leftrightarrow uyv \in L \text{ for all } u, v \in \Sigma^*.$$

This congruence is also known as the *syntactic congruence* of L . The quotient set Σ^+ / \approx_L of equivalence classes of the relation \approx_L is a semigroup called the *syntactic semigroup* of L , and Σ^* / \approx_L is the *syntactic monoid* of L . The *syntactic complexity* $\sigma(L)$ of L is the cardinality of its syntactic semigroup. A language is regular if and only if its syntactic semigroup is finite. We consider only regular languages in this paper; so we assume all syntactic semigroups and syntactic monoids are finite in the following discussion.

A DFA is denoted by $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$, as usual. The DFA \mathcal{A} accepts a word $w \in \Sigma^*$ if $\delta(q_1, w) \in F$. The language accepted by \mathcal{A} is denoted by $L(\mathcal{A})$. If q is a state of \mathcal{A} , then the language L_q of q is the language accepted by the DFA $(Q, \Sigma, \delta, q, F)$. Two states p and q of \mathcal{A} are *equivalent* if $L_p = L_q$. If $L \subseteq \Sigma^*$ is a regular language, then its *quotient DFA* is $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$, where $Q = \{L_w \mid w \in \Sigma^*\}$, $\delta(L_w, a) = L_{wa}$, $q_1 = L_\varepsilon = L$, $F = \{L_w \mid \varepsilon \in L_w\}$. The *quotient complexity* $\kappa(L)$ of L is the number of distinct quotients of L . The quotient DFA of L is the minimal DFA accepting L , and so quotient complexity is the same as state complexity.

If $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$ is a DFA, then its *transition semigroup* [15], denoted by $T_{\mathcal{A}}$, consists of all transformations t_w on Q performed by nonempty words $w \in \Sigma^+$ such that $it_w = \delta(i, w)$ for all $i \in Q$. The syntactic semigroup T_L of a regular language L is isomorphic to the transition semigroup of the quotient DFA \mathcal{A} of L [13]; so we represent elements of T_L by transformations in $T_{\mathcal{A}}$.

On the other hand, given a set $G = \{t_a \mid a \in \Sigma\}$ of transformations of Q , we can define the transition function δ of some DFA \mathcal{A} such that $\delta(i, a) = it_a$ for all $i \in Q$. The transition semigroup of such a DFA is the semigroup generated by G . When the context is clear we simply write $a = t$, where t is a transformation of Q , to mean that the transformation performed by $a \in \Sigma$ is t .

3 \mathcal{R} -Trivial Regular Languages

Given DFA $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$, we can define the *reachability relation* \rightarrow as follows. For all $p, q \in Q$, $p \rightarrow q$ if and only if $\delta(p, w) = q$ for some $w \in \Sigma^*$. We say that \mathcal{A} is *partially ordered* [1] if the relation \rightarrow is a partial order on Q .

Consider the natural order $<$ on Q . A transformation t of Q is *nondecreasing* if $p \leq pt$ for all $p \in Q$. The set \mathcal{F}_Q of all nondecreasing transformations of Q is a semigroup, since the composition of two nondecreasing transformations is again nondecreasing. It was shown in [1] that a language L is \mathcal{R} -trivial if and only if its quotient DFA is partially ordered. Hence, equivalently, L is an \mathcal{R} -trivial language if and only if its syntactic semigroup contains only nondecreasing transformations.

A transformation t of Q is an *idempotent* if $t^2 = t$. It is known [7] that the semigroup \mathcal{F}_Q can be generated by the following set

$$\mathcal{GF}_Q = \{\mathbf{1}_Q\} \cup \{t \in \mathcal{F}_Q \mid t^2 = t \text{ and } \text{rank}(t) = n - 1\}.$$

For any transformation t of Q , let $\text{Fix}(t) = \{i \in Q \mid it = i\}$. Then

Lemma 1. *For any $t \in \mathcal{GF}_Q$, $\text{rng}(t) = \text{Fix}(t)$.*

Proof. Pick arbitrary $t \in \mathcal{GF}_Q$. The claim holds trivially for $\mathbf{1}_Q$. Assume $t \neq \mathbf{1}_Q$. Clearly $\text{Fix}(t) \subseteq \text{rng}(t)$. Suppose there exists $i \in \text{rng}(t)$ but $it \neq i$. Then $jt = i$ for some $j \in Q$, and $j \neq i$. However, since $jt^2 = it \neq i = jt$, t is not an idempotent, which is a contradiction. Therefore $\text{rng}(t) = \text{Fix}(t)$. \square

If $n = 1$, then \mathcal{F}_Q contains only the identity transformation $\mathbf{1}_Q$, and $\mathcal{GF}_Q = \mathcal{F}_Q = \{\mathbf{1}_Q\}$. So $|\mathcal{GF}_Q| = |\mathcal{F}_Q| = 1$. If $n \geq 2$, then we have

Lemma 2. *For $n \geq 2$, $|\mathcal{GF}_Q| = 1 + C_2^n$.*

Proof. Pick $t \in \mathcal{GF}_Q$ such that $t \neq \mathbf{1}_Q$. Then $\text{rank}(t) = n - 1$, and, by Lemma 1, $|\text{Fix}(t)| = n - 1$. There is only one element $i \in Q \setminus \text{Fix}(t)$, and $i < it$. Note that t is fully determined by the pair (i, it) . Hence there are C_2^n different t . Together with the identity $\mathbf{1}_Q$, the cardinality of \mathcal{GF}_Q is $1 + C_2^n$. \square

Lemma 3. *If $G \subseteq \mathcal{F}_Q$ and G generates \mathcal{F}_Q , then $\mathcal{GF}_Q \subseteq G$.*

Proof. Suppose there exists $t \in \mathcal{GF}_Q$ such that $t \notin G$. Since G generates \mathcal{F}_Q , t can be written as $t = g_1 \cdots g_k$ for some $g_1, \dots, g_k \in G$, where $k \geq 2$. Then $\text{rng}(g_1) \supseteq \cdots \supseteq \text{rng}(g_k) \supseteq \text{rng}(t)$. Note that $\mathbf{1}_Q$ is the only element in \mathcal{F}_Q with $\text{rank}(t) = n - 1$; so if $t = \mathbf{1}_Q$, then $g_1 = \cdots = g_k = \mathbf{1}_Q$, a contradiction.

Assume $t \neq \mathbf{1}_Q$. Then $\text{rank}(t) = n - 1$, and $\text{rng}(g_1) = \cdots = \text{rng}(g_k) = \text{rng}(t)$. Since each g_i is nondecreasing, for all $p \in \text{Fix}(t)$, we must have $p \in \text{Fix}(g_i)$ as well; so $\text{Fix}(t) \subseteq \text{Fix}(g_i)$. Moreover, since $\text{Fix}(g_i) \subseteq \text{rng}(g_i) = \text{rng}(t)$ and $\text{rng}(t) = \text{Fix}(t)$ by Lemma 1, $\text{Fix}(g_i) = \text{Fix}(t) = \text{rng}(t)$. Now, let q be the unique element in $Q \setminus \text{Fix}(t)$. Then $q \notin \text{Fix}(g_1)$, and $qg_1 \in \text{Fix}(g_2) = \cdots = \text{Fix}(g_k)$. So $q(g_1 \cdots g_k) = qg_1$. However, since $t = g_1 \cdots g_k$, $q(g_1 \cdots g_k) = qt$ and $qg_1 = qt$. Hence $g_1 = t$, and we get a contradiction again. Therefore $\mathcal{GF}_Q \subseteq G$. \square

Consequently, \mathcal{GF}_Q is the unique minimal generator of \mathcal{F}_Q . So we obtain

Theorem 1. *If $L \subseteq \Sigma^*$ is a regular \mathcal{R} -trivial language of quotient complexity $\kappa(L) = n \geq 1$, then its syntactic complexity $\sigma(L)$ satisfies $\sigma(L) \leq n!$, and this bound is tight if $|\Sigma| = 1$ for $n = 1$ and $|\Sigma| \geq 1 + C_2^n$ for $n \geq 2$.*

Proof. Let \mathcal{A} be the quotient DFA of L , and let T_L be its syntactic semigroup. Then T_L is a subset of \mathcal{F}_Q . Pick an arbitrary $t \in \mathcal{F}_Q$. For each $p \in Q$, since $p \leq pt$, pt can be chosen from $\{p, p+1, \dots, n\}$. Hence there are exactly $n!$ transformations in \mathcal{F}_Q , and $\sigma(L) \leq n!$.

When $n = 1$, the only regular languages are ε or \emptyset , and they both are \mathcal{R} -trivial. To see the bound is tight for $n \geq 2$, let $\mathcal{A}_n = (Q, \Sigma, \delta, 1, \{n\})$ be the DFA with alphabet Σ of size $1 + C_2^n$ and set of states $Q = \{1, \dots, n\}$, where each $a \in \Sigma$ defines a distinct transformation in \mathcal{GF}_Q . For each $p \in Q$, since \mathcal{GF}_Q generates \mathcal{F}_Q and $t_p = [p, n, \dots, n] \in \mathcal{F}_Q$, $t_p = e_1 \cdots e_k$ for some $e_1, \dots, e_k \in \mathcal{GF}_Q$, where k depends on p . Then there exist $a_1, \dots, a_k \in \Sigma$ such that each a_i performs e_i and state p is reached by $w = a_1 \dots a_k$. Moreover, since $t = [2, 3, \dots, n, n] \in \mathcal{F}_Q$, there exist $b_1, \dots, b_l \in \Sigma$ such that the word $u = b_1 \dots b_l$ performs t . So state $p \in Q$ can be distinguished from other states by the word u^{n-p} . Hence $L = L(\mathcal{A}_n)$ has quotient complexity $\kappa(L) = n$. The syntactic monoid of L is \mathcal{F}_Q , and so $\sigma(L) = n!$. By Lemma 3, the alphabet of \mathcal{A}_n is minimal. \square

Example 1. When $n = 4$, there are $4! = 24$ nondecreasing transformations of $Q = \{1, 2, 3, 4\}$. Among them, there are 11 transformations with rank $n - 1 = 3$. The following 6 transformations from the 11 are idempotents:

$$\begin{aligned} e_1 &= [1, 2, 4, 4], & e_2 &= [1, 3, 3, 4] \\ e_3 &= [1, 4, 3, 4], & e_4 &= [2, 2, 3, 4] \\ e_5 &= [3, 2, 3, 4], & e_6 &= [4, 2, 3, 4] \end{aligned}$$

Together with the identity transformation $\mathbf{1}_Q$, we have the generating set \mathcal{GF}_Q for \mathcal{F}_Q with 7 transformations. We can then define the DFA \mathcal{A}_4 with 7 inputs as in the proof of Theorem 1; \mathcal{A}_4 is shown in Fig. 1. The quotient complexity of $L = L(\mathcal{A}_4)$ is 4, and the syntactic complexity of L is 24. \blacksquare

4 \mathcal{J} -Trivial Regular Languages

We first recall some facts from universal algebra. Let Q be a nonempty finite set with n elements, and assume without loss of generality that $Q = \{1, 2, \dots, n\}$. There is a linear order on Q , namely the natural order $<$ on integers. If X is a nonempty subset of Q , then the maximal element in X is denoted by $\max(X)$. A *partition* π of Q is a collection $\pi = \{X_1, X_2, \dots, X_m\}$ of nonempty subsets of Q such that

1. $Q = X_1 \cup X_2 \cup \dots \cup X_m$, and

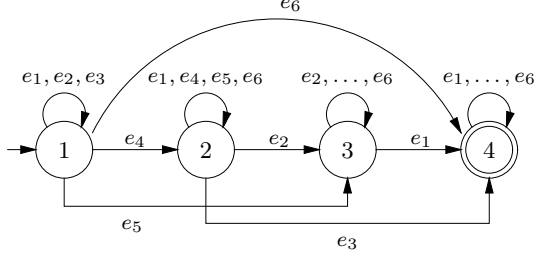


Fig. 1. DFA \mathcal{A}_4 with $\kappa(L(\mathcal{A}_4)) = 4$ and $\sigma(L(\mathcal{A}_4)) = 24$; the input performing the identity transformation is not shown.

2. $X_i \cap X_j = \emptyset$ for all $1 \leq i < j \leq m$.

We call each subset X_i a *block* in π . For any partition π of Q , let $\text{Max}(\pi) = \{\max(X) \mid X \in \pi\}$. The set of all partitions of Q is denoted by Π_Q . We can define a partial order \preceq on Π_Q such that, for any $\pi_1, \pi_2 \in \Pi_Q$, $\pi_1 \preceq \pi_2$ if and only if each block of π_1 is contained in some block of π_2 . We say π_1 *refines* π_2 if $\pi_1 \preceq \pi_2$. Then (Π_Q, \preceq) forms a poset. Furthermore, (Π_Q, \preceq) is a finite lattice; for any $\pi_1, \pi_2 \in \Pi_Q$, their *meet* $\pi_1 \wedge \pi_2$ is the \preceq -largest partition that refines both π_1 and π_2 , and their *join* $\pi_1 \vee \pi_2$ is the \preceq -smallest partition that is refined by both π_1 and π_2 . From now on, we simply refer to the lattice (Π_Q, \preceq) as Π_Q .

For any $m \geq 1$, we can define an equivalence relation \leftrightarrow_m on Σ^* as follows. For any $u, v \in \Sigma^*$, $u \leftrightarrow_m v$ if and only if for every $x \in \Sigma^*$ with $|x| \leq m$,

$$x \text{ is a subword of } u \Leftrightarrow x \text{ is a subword of } v.$$

Let L be any language over Σ . Then L is *piecewise-testable* if there exists $m \geq 1$ such that, for every $u, v \in \Sigma^*$, $u \leftrightarrow_m v$ implies that $u \in L \Leftrightarrow v \in L$. Let $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$ be a DFA. If Γ is a subset of Σ , a *component* of \mathcal{A} restricted to Γ is a minimal subset P of Q such that, for all $p \in Q$ and $w \in \Gamma^*$, $\delta(p, w) \in P$ if and only if $p \in P$. A state q of \mathcal{A} is *maximal* if $\delta(q, a) = q$ for all $a \in \Sigma$. Simon [19] proved the following characterization of piecewise-testable languages.

Theorem 2 (Simon). *Let L be a regular language over Σ , let \mathcal{A} be its quotient DFA, and let T_L be its syntactic monoid. Then the following are equivalent:*

1. L is piecewise-testable;
2. \mathcal{A} is partially ordered, and for every nonempty subset Γ of Σ , each component of \mathcal{A} restricted to Γ has exactly one maximal state;
3. T_L is \mathcal{J} -trivial.

Consequently, a regular language is piecewise-testable if and only if it is \mathcal{J} -trivial. The following theorem is due to Saito [16]. It is another characterization of \mathcal{J} -trivial monoids.

Theorem 3 (Saito). *Let S be a monoid of transformations of Q . Then the following are equivalent:*

1. S is \mathcal{J} -trivial;
2. S is a subset of \mathcal{F}_Q and $\Omega(ts) = \Omega(t) \vee \Omega(s)$ for all $t, s \in S$.

Let L be a regular \mathcal{J} -trivial language with quotient DFA $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$ and syntactic monoid T_L . Since T_L is a subset of \mathcal{F}_Q , to get an upper bound on the syntactic complexity of L , we find an upper bound on the cardinality of \mathcal{J} -trivial submonoids of \mathcal{F}_Q .

Lemma 4. *If $t, s \in \mathcal{F}_Q$, then*

1. $\text{Fix}(t) = \text{Max}(\Omega(t))$;
2. $\Omega(t) \preceq \Omega(s)$ implies that $\text{Fix}(t) \supseteq \text{Fix}(s)$, where the equality holds if and only if $\Omega(t) = \Omega(s)$;

Proof. 1. First, for each $j \in \text{Max}(\Omega(t))$, since $t \in \mathcal{F}_Q$, we have $jt = j$, and $j \in \text{Fix}(t)$. So $\text{Max}(\Omega(t)) \subseteq \text{Fix}(t)$. On the other hand, if there exists $j \in \text{Fix}(t) \setminus \text{Max}(\Omega(t))$, then $jt = j$, and $j < \max(\omega_t(j))$. Let $i = \max(\omega_t(j))$; then for any $k, l \geq 0$, $jt^k = j < i = it^l$. So $i \notin \omega_t(j)$, which is a contradiction. Hence $\text{Fix}(t) = \text{Max}(\Omega(t))$.

2. Assume $\Omega(t) \preceq \Omega(s)$. By definition, we have $\text{Max}(\Omega(t)) \supseteq \text{Max}(\Omega(s))$. Then, by 1, $\text{Fix}(t) \supseteq \text{Fix}(s)$. Furthermore, $\Omega(t) = \Omega(s)$ if and only if $\text{Max}(\Omega(t)) = \text{Max}(\Omega(s))$, and if and only if $\text{Fix}(t) = \text{Fix}(s)$. \square

Example 2. Consider nondecreasing $t = [1, 3, 3, 5, 6, 6]$, as shown in Fig. 2 (a). The orbit set $\Omega(t)$ has three blocks: $\{1\}$, $\{2, 3\}$, and $\{4, 5, 6\}$. Note that $\text{Fix}(t) = \{1, 3, 6\} = \text{Max}(\Omega(t))$, as expected.

In addition, let $s = [4, 3, 3, 6, 6, 6]$ be another nondecreasing transformation, as shown in Fig. 2 (b). The orbit set $\Omega(s)$ has two blocks: $\{1, 4, 5, 6\}$ and $\{2, 3\}$. Note that $\Omega(t) \prec \Omega(s)$ and $\text{Fix}(t) \supset \text{Fix}(s)$. \blacksquare

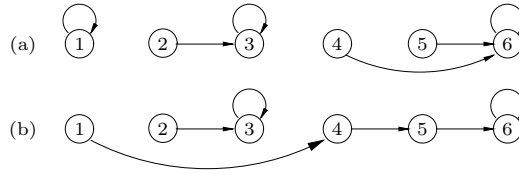


Fig. 2. Nondecreasing transformations $t = [1, 3, 3, 5, 6, 6]$ and $s = [4, 3, 3, 6, 6, 6]$.

Define the transformation $t_{\max} = [2, 3, \dots, n, n]$. The subscript “max” is chosen because $\Omega(t_{\max}) = \{Q\}$ is the maximum element in the lattice Π_Q . Clearly $t_{\max} \in \mathcal{F}_Q$ and $\text{Fix}(t_{\max}) = \{n\}$. For any submonoid S of \mathcal{F}_Q , let $S[t_{\max}]$ be the smallest monoid containing t_{\max} and all elements of S .

Lemma 5. *Let S be a \mathcal{J} -trivial submonoid of \mathcal{F}_Q . Then*

1. $S[t_{\max}]$ is \mathcal{J} -trivial.
2. Let $\mathcal{A} = (Q, \Sigma, \delta, 1, \{n\})$ be the DFA in which each $a \in \Sigma$ defines a distinct transformation in $S[t_{\max}]$. Then \mathcal{A} is minimal.

Proof. 1. By Theorem 3, it is sufficient to prove that for any $t \in S$, $\Omega(t) \vee \Omega(t_{\max}) = \Omega(tt_{\max})$ and $\Omega(t_{\max}) \vee \Omega(t) = \Omega(t_{\max}t)$. Note that $\Omega(t_{\max}) = \{Q\}$; so we have $\Omega(t) \vee \Omega(t_{\max}) = \Omega(t_{\max}) \vee \Omega(t) = \{Q\}$. On the other hand, since $S \subseteq \mathcal{F}_Q$ and $t_{\max} \in \mathcal{F}_Q$, both tt_{\max} and $t_{\max}t$ are nondecreasing as well. Suppose $i \in \text{Fix}(tt_{\max})$; then $i(tt_{\max}) = (it)t_{\max} = i$. Since t_{\max} is nondecreasing, $it \leq i$; and since t is also nondecreasing, $i \leq it$. Hence $it = i$, and $it_{\max} = i$, which implies that $i \in \text{Fix}(t_{\max})$ and $i = n$. Then $\text{Fix}(tt_{\max}) = \{n\}$ and $\Omega(tt_{\max}) = \{Q\}$. Similarly, $\text{Fix}(t_{\max}t) = \{n\}$ and $\Omega(t_{\max}t) = \{Q\}$. Therefore $S[t_{\max}]$ is also \mathcal{J} -trivial.

2. Suppose $a_0 \in \Sigma$ performs the transformation t_{\max} . Each state $p \in Q$ can be reached from the initial state 1 by the word $u = a_0^{p-1}$, and p accepts the word $v = a_0^{n-p}$, while all other states reject v . So \mathcal{A} is minimal. \square

For any \mathcal{J} -trivial submonoid S of \mathcal{F}_Q , we denote by $\mathcal{A}(S, t_{\max})$ the DFA in Lemma 5. Then $\mathcal{A}(S, t_{\max})$ is the quotient DFA of some regular \mathcal{J} -trivial language L . Next, we have

Lemma 6. *Let S be a \mathcal{J} -trivial submonoid of \mathcal{F}_Q . For any $t, s \in S$, if $\text{Fix}(t) = \text{Fix}(s)$, then $\Omega(t) = \Omega(s)$.*

Proof. Pick any $t, s \in S$ such that $\text{Fix}(t) = \text{Fix}(s)$. If $t = s$, then it is trivial that $\Omega(t) = \Omega(s)$. Assume $t \neq s$, and $\Omega(t) \neq \Omega(s)$. By Part 2 of Lemma 4, we have $\Omega(t) \not\leq \Omega(s)$ and $\Omega(s) \not\leq \Omega(t)$. Then there exists $i \in Q$ such that $\omega_t(i) \neq \omega_s(i)$. Suppose $p = \max(\omega_t(i))$ and $q = \max(\omega_s(i))$; then $p, q \in \text{Fix}(t) = \text{Fix}(s)$, and $p \neq q$. Consider the DFA $\mathcal{A}(S, t_{\max})$ with alphabet Σ , and suppose that $a \in \Sigma$ performs t and $b \in \Sigma$ performs s . Let \mathcal{B} be the DFA $\mathcal{A}(S, t_{\max})$ restricted to $\{a, b\}$. Since $p \in \omega_t(i)$ and $q \in \omega_s(i)$, then p, q are in the same component P of \mathcal{B} . However, p and q are two distinct maximal states in P , which contradicts Theorem 2. Therefore $\Omega(t) = \Omega(s)$. \square

Example 3. To illustrate one usage of Lemma 6, we consider two nondecreasing transformations $t = [2, 2, 4, 4]$ and $s = [3, 2, 4, 4]$. They have the same set of fixed points $\text{Fix}(t) = \text{Fix}(s) = \{2, 4\}$. However, $\Omega(t) = \{\{1, 2\}, \{3, 4\}\}$ and $\Omega(s) = \{\{2\}, \{1, 3, 4\}\}$. By Lemma 6, t and s cannot appear together in a \mathcal{J} -trivial monoid. Indeed, consider any minimal DFA \mathcal{A} having at least two inputs a, b such that a performs t and b performs s . The DFA \mathcal{B} of \mathcal{A} restricted to the alphabet $\{a, b\}$ is shown in Fig. 3. There is only one component in \mathcal{B} , but there are two maximal states 2 and 4. By Theorem 2, the syntactic monoid of \mathcal{A} is not \mathcal{J} -trivial. \blacksquare

For any partition π of Q , define $\mathcal{E}(\pi) = \{t \in \mathcal{F}_Q \mid \Omega(t) = \pi\}$. Then

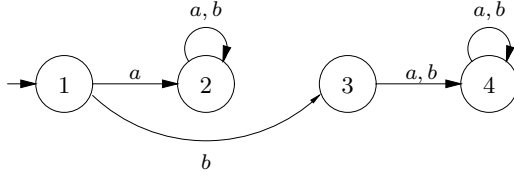


Fig. 3. DFA \mathcal{B} with two inputs a and b , where $t_a = [2, 2, 4, 4]$ and $t_b = [3, 2, 4, 4]$.

Lemma 7. *If π is a partition of Q with r blocks, where $1 \leq r \leq n$, then $|\mathcal{E}(\pi)| \leq (n - r)!$.*

Proof. Suppose $\pi = \{X_1, \dots, X_r\}$, and $|X_i| = k_i$ for each i , $1 \leq i \leq r$. Without loss of generality, we can rearrange subsets X_i 's such that $k_1 \leq \dots \leq k_r$. Let $t \in \mathcal{E}(\pi)$ be any transformation. Then $t \in \mathcal{F}_Q$, and hence $\text{Fix}(t) = \text{Max}(\Omega(t)) = \text{Max}(\pi)$. Consider each block X_i , and suppose $X_i = \{j_1, \dots, j_{k_i}\}$ such that $j_1 < \dots < j_{k_i}$. Since $j_{k_i} = \max(X_i)$, then $j_{k_i} \in \text{Fix}(t)$ and $j_{k_i}t = j_{k_i}$. On the other hand, if $1 \leq l < k_i$, then $j_l \notin \text{Max}(\pi)$, and since $t \in \mathcal{F}_Q$, we have $j_lt > j_l$; since $j_lt \in \omega_t(j_l) = X_i$, $j_lt \in \{j_{l+1}, \dots, j_{k_i}\}$. So there are $(k_i - 1)!$ different $t|_{X_i}$, and there are $\prod_{i=1}^r (k_i - 1)!$ different transformations t in $\mathcal{E}(\pi)$.

Clearly, if $r = 1$, then $k_r = n$ and $|\mathcal{E}(\pi)| = (n - 1)!$. Assume $r \geq 2$. Note that $k_i \geq 1$ for all i , $1 \leq i \leq r$, and $\sum_{i=1}^r k_i = n$. If $k_1 = \dots = k_{r-1} = 1$, then $k_r = n - r + 1$, and $|\mathcal{E}(\pi)| = (k_r - 1)! \prod_{i=1}^{r-1} 0! = (n - r)!$. Otherwise, let h be the smallest index such that $k_h > 1$. Then

$$\begin{aligned} \prod_{i=1}^r (k_i - 1)! &= \prod_{i=1}^{h-1} 0! \prod_{i=h}^r (k_i - 1)! \\ &= (k_r - 1)! \prod_{i=h}^{r-1} (k_i - 1)! \end{aligned}$$

Since $(k_h - 1)! < (k_h - 1)^{k_h - 1} \leq (k_r - 1)^{k_h - 1} < (k_r + k_h - 2) \cdots k_r$:

$$< (k_r + k_h - 2)! \prod_{i=h+1}^{r-1} (k_i - 1)!$$

Similarly, we have that

$$\begin{aligned} &< (k_r + k_h + \dots + k_{r-1} - (r - h + 1))! \\ &= (n - r)! \end{aligned}$$

Therefore $|\mathcal{E}(\pi)| \leq (n - r)!$. □

Example 4. Suppose $n = 10$, $r = 3$, and consider the partition $\pi = \{X_1, X_2, X_3\}$, where $X_1 = \{1, 2, 5\}$, $X_2 = \{3, 7\}$, and $X_3 = \{4, 6, 8, 9, 10\}$. Then $k_1 = |X_1| = 3$, $k_2 = |X_2| = 2$, and $k_3 = |X_3| = 5$. Let $t \in \mathcal{E}(\pi)$ be an arbitrary transformation; then $\text{Fix}(t) = \{5, 7, 10\}$. For any $i \in X_1$, if $i = 1$, then it could be 2 or 5; otherwise $i = 2$ or 5, and it must be 5. So there are $(k_1 - 1)! = 2!$ different $t|_{X_1}$. Similarly, there are $(k_2 - 1)! = 1!$ different $t|_{X_2}$ and $(k_3 - 1)! = 4!$ different $t|_{X_3}$. Hence we have $|\mathcal{E}(\pi)| = 2!1!4! = 48$.

Consider another partition $\pi' = \{X'_1, X'_2, X'_3\}$ with three blocks, where $X'_1 = \{5\}$, $X'_2 = \{7\}$, and $X'_3 = \{1, 2, 3, 4, 6, 8, 9, 10\}$. Then $k_1 = |X'_1| = 1$, $k_2 = |X'_2| = 1$, and $k_3 = |X'_3| = 8$. We have that $\text{Max}(\pi') = \text{Max}(\pi) = \{5, 7, 10\}$. Then, for any $t \in \mathcal{E}(\pi')$, $\text{Fix}(t) = \{5, 7, 10\}$ as well. Since $k_1 = k_2 = 1$, both $t|_{X_1}$ and $t|_{X_2}$ are unique. There are $(k_3 - 1)! = 7!$ different $t|_{X_3}$. Together we have $|\mathcal{E}(\pi')| = 1!1!7! = (10 - 3)! = 5040$, which is the upper bound in Lemma 7 for $n = 10$ and $r = 3$. ■

Note that, for any $t \in \mathcal{F}_Q$, we have $n \in \text{Fix}(t)$. Let $\mathcal{P}_n(Q)$ be the set of all subsets Z of Q such that $n \in Z$. Then we obtain the following upper bound.

Proposition 1. *If S is a \mathcal{J} -trivial submonoid of \mathcal{F}_Q , then*

$$|S| \leq \sum_{r=1}^n C_{r-1}^{n-1} (n-r)! = \lfloor e(n-1)! \rfloor.$$

Proof. Assume S is a \mathcal{J} -trivial submonoid of \mathcal{F}_Q . For any $Z \in \mathcal{P}_n(Q)$, let $S_Z = \{t \in S \mid \text{Fix}(t) = Z\}$. Then $S = \bigcup_{Z \in \mathcal{P}_n(Q)} S_Z$, and for any $Z_1, Z_2 \in \mathcal{P}_n(Q)$ with $Z_1 \neq Z_2$, $S_{Z_1} \cap S_{Z_2} = \emptyset$.

Pick any $Z \in \mathcal{P}_n(Q)$. By Lemma 6, for any $t, s \in S_Z$, since $\text{Fix}(t) = \text{Fix}(s) = Z$, we have $\Omega(t) = \Omega(s)$. Let $\pi \in \Pi_Q$ denote such a partition $\Omega(t)$ of Q . Suppose $r = |Z|$. Since $n \in Z$, we have $r \geq 1$; and clearly $r \leq n$. Note that $S_Z \subseteq \mathcal{E}(\pi)$. By Lemma 7, $|S_Z| \leq |\mathcal{E}(\pi)| = (n-r)!$. Since there are C_{r-1}^{n-1} different Z , we have that

$$\begin{aligned} |S| &= \sum_{Z \in \mathcal{P}_n(Q)} |S_Z| \leq \sum_{r=1}^n C_{r-1}^{n-1} (n-r)! \\ &= \sum_{r=1}^n \frac{(n-1)!}{(r-1)!} \\ &= \lfloor e(n-1)! \rfloor. \end{aligned}$$

The last equality is due to a well-known combinatorics identity. □

The above upper bound is met by the following monoid \mathcal{S}_n . For any $Z \in \mathcal{P}_n(Q)$, suppose $Z = \{j_1, \dots, j_r\}$ such that $j_1 < \dots < j_r$; then we define partition $\pi_Z = \{Q\}$ if $Z = \{n\}$, and $\pi_Z = \{\{j_1\}, \dots, \{j_{r-1}\}, Q \setminus \{j_1, \dots, j_{r-1}\}\}$ otherwise. Let

$$\mathcal{S}_n = \bigcup_{Z \in \mathcal{P}_n(Q)} \mathcal{E}(\pi_Z).$$

Example 5. Suppose $n = 4$; then $|\mathcal{P}_4(Q)| = 2^3 = 8$. First consider $Z = \{1, 3, 4\} \in \mathcal{P}_4(Q)$. By definition, $\pi_Z = \{\{1\}, \{3\}, \{2, 4\}\}$. There is only one transformation $t_1 = [1, 4, 3, 4]$ in $\mathcal{E}(\pi_Z)$. If $Z' = \{3, 4\}$, then $\pi_{Z'} = \{\{3\}, \{1, 2, 4\}\}$. There are two transformations $t_2 = [2, 4, 3, 4]$ and $t_3 = [4, 4, 3, 4]$ in $\mathcal{E}(\pi_{Z'})$. Table 1 summaries the number of transformations in $\mathcal{E}(\pi_Z)$ for each $Z \in \mathcal{P}_4(Q)$. Note that the set \mathcal{S}_4 contains 16 transformations in total. ■

Table 1. Number of transformations in $\mathcal{E}(\pi_Z)$ for each $Z \in \mathcal{P}_4(Q)$.

Z	Blocks of π_Z	$ \mathcal{E}(\pi_Z) $
$\{1, 2, 3, 4\}$	$\{1\}, \{2\}, \{3\}, \{4\}$	1
$\{1, 2, 4\}$	$\{1\}, \{2\}, \{3, 4\}$	1
$\{1, 3, 4\}$	$\{1\}, \{3\}, \{2, 4\}$	1
$\{2, 3, 4\}$	$\{2\}, \{3\}, \{1, 4\}$	1
$\{1, 4\}$	$\{1\}, \{2, 3, 4\}$	2
$\{2, 4\}$	$\{2\}, \{1, 3, 4\}$	2
$\{3, 4\}$	$\{3\}, \{1, 2, 4\}$	2
$\{4\}$	$\{1, 2, 3, 4\}$	6

Proposition 2. The set \mathcal{S}_n is a \mathcal{J} -trivial submonoid of \mathcal{F}_Q with cardinality

$$g(n) = |\mathcal{S}_n| = \sum_{r=1}^n C_{r-1}^{n-1} (n-r)! = \lfloor e(n-1)! \rfloor. \quad (1)$$

Proof. First we prove the following claim:

Claim: For any $t, s \in \mathcal{S}_n$, $\Omega(ts) = \pi_Z$ for some $Z \in \mathcal{P}_n(Q)$.

Let $t \in \mathcal{E}(\pi_{Z_1})$ and $s \in \mathcal{E}(\pi_{Z_2})$ for some $Z_1, Z_2 \in \mathcal{P}_n(Q)$. Suppose $\Omega(ts) \neq \pi_Z$ for any $Z \in \mathcal{P}_n(Q)$. Then there exists a block $X_0 \in \Omega(ts)$ such that $n \notin X_0$ and $|X_0| \geq 2$. Suppose $i \in X_0$ with $i \neq \max(X_0)$. We must have $i \in \omega_t(n)$ or $it \in \omega_s(n)$; otherwise $it = i$ and $(it)s = i$ and so $i = \max(X_0)$. However, in either case, there exists large m such that $it^m = n$ or $(it)s^m = n$, respectively. Then $n \in \omega_{ts}(i) = X_0$, a contradiction. So the claim holds. ■

By the claim, for any $t, s \in \mathcal{S}_n$, since $\Omega(ts) = \pi_Z$ for some $Z \in \mathcal{P}_n(Q)$, $ts \in \mathcal{E}(\pi_Z) \subseteq \mathcal{S}_n$. Hence \mathcal{S}_n is a submonoid of \mathcal{F}_Q .

Next we show that \mathcal{S}_n is \mathcal{J} -trivial. Pick any $t, s \in \mathcal{S}_n$, and suppose $t \in \mathcal{E}(\pi_{Z_1})$ and $s \in \mathcal{E}(\pi_{Z_2})$ for some $Z_1, Z_2 \in \mathcal{P}_n(Q)$. Suppose $\text{Max}(Z_1) \cap \text{Max}(Z_2) = \{j_1, \dots, j_r\}$, for some $r \geq 0$. Then we have $Z_1 \vee Z_2 = \{\{j_1\}, \dots, \{j_r\}, X\}$, where $X = Q \setminus \{j_1, \dots, j_r\}$ and $n \in X$. On the other hand, by the claim, $\Omega(ts) = \{\{p_1\}, \dots, \{p_k\}, Y\}$, where $Y = Q \setminus \{p_1, \dots, p_k\}$. Note that, since $\mathcal{E}(\pi_{Z_1}), \mathcal{E}(\pi_{Z_2}) \subseteq \mathcal{F}_Q$, $\text{Max}(\Omega(ts)) = \text{Fix}(ts) = \text{Fix}(t) \cap \text{Fix}(s) = \text{Max}(Z_1) \cap \text{Max}(Z_2)$. Then $r = k$

and $\{j_1, \dots, j_r\} = \{p_1, \dots, p_k\}$. Hence $\Omega(t) \vee \Omega(s) = Z_1 \vee Z_2 = \Omega(ts)$. By Theorem 3, \mathcal{S}_n is \mathcal{J} -trivial.

For any $Z \in \mathcal{P}_n(Q)$ with $|Z| = r$, where $1 \leq r \leq n$, suppose $\pi_Z = \{X_1, \dots, X_r\}$ with $k_i = |X_i| = 1$ for $1 \leq i < r$, and $k_r = |X_r|$. By Lemma 7, $|\mathcal{E}(\pi_Z)| = (n-r)!$. Since $n \in Z$ is fixed, there are C_{r-1}^{n-1} different Z . Therefore $|\mathcal{S}_n| = \sum_{r=1}^n C_{r-1}^{n-1} (n-r)! = \lfloor e(n-1)! \rfloor$. \square

Let t be any transformation of Q . An orbit X of t is *trivial* if it contains just one element of Q ; otherwise it is *non-trivial*. Hence any transformation $t \in \mathcal{S}_n$ has only one non-trivial orbit. We now define a generating set of the monoid \mathcal{S}_n .

Definition 1. Suppose $n \geq 1$. For any $Z \in \mathcal{P}_n(Q)$, if $Z = Q$, then let $t_Z = \mathbf{1}_Q$. Otherwise, let $h_Z = \max(Q \setminus Z)$, and let t_Z be a transformation of Q defined by: For all $i \in Q$,

$$it \stackrel{\text{def}}{=} \begin{cases} i & \text{if } i \in Z, \\ n & \text{if } i = h_Z, \\ h_Z & \text{otherwise.} \end{cases}$$

Let $\mathcal{GS}_n = \{t_Z \mid Z \in \mathcal{P}_n(Q)\}$.

Example 6. Suppose $n = 5$. As the first example, consider $Z = \{1, 3, 4, 5\}$. Then $h_Z = \max(Q \setminus Z) = 2$, and $t_Z = [1, 5, 3, 4, 5]$. If $Z' = \{4, 5\}$, then $h_{Z'} = 4$ and $t_{Z'} = [3, 3, 5, 4, 5]$. If $Z'' = \{5\}$, then $h_{Z''} = 4$ and $t_{Z''} = [4, 4, 4, 5, 5]$. The set \mathcal{GS}_5 contains the following 16 transformations:

$$\begin{array}{lll} t_1 = [1, 2, 3, 4, 5], & t_2 = [1, 2, 3, 5, 5], & t_3 = [1, 2, 4, 5, 5], \\ t_4 = [1, 2, 5, 4, 5], & t_5 = [1, 3, 5, 4, 5], & t_6 = [1, 4, 3, 5, 5], \\ t_7 = [1, 4, 4, 5, 5], & t_8 = [1, 5, 3, 4, 5], & t_9 = [2, 5, 3, 4, 5], \\ t_{10} = [3, 2, 5, 4, 5], & t_{11} = [3, 3, 5, 4, 5], & t_{12} = [4, 2, 3, 5, 5], \\ t_{13} = [4, 2, 4, 5, 5], & t_{14} = [4, 4, 3, 5, 5], & t_{15} = [4, 4, 4, 5, 5], \\ t_{16} = [5, 2, 3, 4, 5]. \end{array}$$

■

Proposition 3. For $n \geq 1$, the monoid \mathcal{S}_n can be generated by the set \mathcal{GS}_n of 2^{n-1} transformations of Q .

Proof. First, for any $t_Z \in \mathcal{GS}_n$, where $Z \in \mathcal{P}_n(Q)$, we have $\Omega(t_Z) = \pi_Z$; hence $t_Z \in \mathcal{E}(\pi_Z) \subseteq \mathcal{S}_n$. So $\mathcal{GS}_n \subseteq \mathcal{S}_n$ and $\langle \mathcal{GS}_n \rangle \subseteq \mathcal{S}_n$.

Fix arbitrary $Z \in \mathcal{P}_n(Q)$, and suppose $U = Q \setminus Z$. Note that $n \in Z$. Let Y be the block in π_Z such that $n \in Y$. For any $t \in \mathcal{E}(\pi_Z)$, we have $\text{Fix}(t) = Z$. Furthermore, if $i \in Q \setminus Y$, then $i \in \text{Fix}(t)$ and $it = i$. We prove by induction on $|U|$ that $\mathcal{E}(\pi_Z) \subseteq \langle \mathcal{GS}_n \rangle$.

1. $U = \emptyset$: Then $\pi_Z = \{\{1\}, \dots, \{n\}\}$, and $\mathcal{E}(\pi_Z) = \{\mathbf{1}_Q\}$. Note that $\mathbf{1}_Q \in \mathcal{GS}_n$. So $\mathcal{E}(\pi_Z) \subseteq \langle \mathcal{GS}_n \rangle$.
2. $U = \{h\}$ for some $h \neq n$: Then $Y = \{h, n\}$. For any $t \in \mathcal{E}(\pi_Z)$, since $\text{Fix}(t) = Z$ and $h \notin Z$, we have $ht > h$. Since Y is an orbit of t , we have $ht = n$, and $t = \binom{h}{n}$. Note that $t_Z = \binom{h}{n} = t$. So $\mathcal{E}(\pi_Z) \subseteq \langle \mathcal{GS}_n \rangle$.

3. $U = \{h_1, h_2\}$ for some $h_1 < h_2 < n$: Then $Y = \{h_1, h_2, n\}$. Note that $t_Z = \binom{h_2}{n} \binom{h_1}{h_2}$. For any $t \in \mathcal{E}(\pi_Z)$, since $h_1 < h_2$, and Y is an orbit of t , we have $h_2 t = n$ and $h_1 t \in \{h_2, n\}$. If $h_1 t = h_2$, then $t = t_Z$; otherwise, $t = t_Z^2$. So $\mathcal{E}(\pi_Z) \subseteq \langle \mathcal{GS}_n \rangle$.
4. $U = \{h_1, \dots, h_l\}$ for some $h_1 < \dots < h_l < n$, where $l \geq 3$: Assume that, for any $Z' \in \mathcal{P}_n(Q)$ with $|Q \setminus Z'| < l$, we have $\mathcal{E}(\pi_{Z'}) \subseteq \langle \mathcal{GS}_n \rangle$. Now $Y = \{h_1, \dots, h_l, n\}$, and $t_Z = \binom{h_l}{n} \binom{h_{l-1}}{h_l} \dots \binom{h_1}{h_l}$. For any $t \in \mathcal{E}(\pi_Z)$, since Y is an orbit of t and $Q \setminus Y \subseteq \text{Fix}(t)$, we have

$$t = \binom{h_l}{p_l} \binom{h_{l-1}}{p_{l-1}} \dots \binom{h_1}{p_1},$$

where $p_l = n$, and $p_i \in \{h_{i+1}, \dots, h_l, n\}$ for $i = 2, \dots, l-1$. We have three cases:

- (a) $p_1 = \dots = p_l = n$: Then $t = t_Z^2$, and $t \in \langle \mathcal{GS}_n \rangle$.
- (b) $p_1 = \dots = p_{l-1} = h_l$: Then $t = t_Z$, and $t \in \langle \mathcal{GS}_n \rangle$ as well.
- (c) Otherwise, there exists some h_i , where $1 \leq i < l$, such that $p_i = h_i t \notin \{h_l, n\}$. Let h_r be the smallest such h_i , and let $Y' = Y \setminus \{h_r\}$. Then $h_r \notin \text{rng}(t)$, and $p_r = h_r t \in Y' \setminus \{h_l, n\}$. Now, let

$$t' = \binom{h_l}{n} \dots \binom{h_{r+1}}{p_{r+1}} \binom{h_{r-1}}{p_{r-1}} \dots \binom{h_1}{p_1}$$

and $Z' = \text{Fix}(t')$. Then Y' is an orbit of t' , and $Z' = \text{Fix}(t) \cup \{h_r\}$; so $t' \in \mathcal{E}(\pi_{Z'})$. By assumption, since $|Q \setminus Z'| = l-1 < l$, we have $t' \in \langle \mathcal{GS}_n \rangle$. As the last step, let $Z'' = \{h_r, p_r\}$. Since $p_r = h_r t > h_r$, we have $t_{Z''} = \binom{p_r}{n} \binom{h_r}{p_r} \in \mathcal{GS}_n$. Note that $h_r \notin \text{rng}(t')$ and $p_r \notin \text{Fix}(t')$. So

$$\begin{aligned} t' t_{Z''} &= \binom{h_l}{n} \dots \binom{h_{r+1}}{p_{r+1}} \binom{h_{r-1}}{p_{r-1}} \dots \binom{h_1}{p_1} \circ \binom{p_r}{n} \binom{h_r}{p_r} \\ &= \binom{h_l}{n} \dots \binom{h_{r+1}}{p_{r+1}} \binom{h_{r-1}}{p_{r-1}} \dots \binom{h_1}{p_1} \circ \binom{h_r}{p_r} \\ &= \binom{h_l}{n} \dots \binom{h_{r+1}}{p_{r+1}} \binom{h_r}{p_r} \binom{h_{r-1}}{p_{r-1}} \dots \binom{h_1}{p_1} \\ &= t. \end{aligned}$$

Thus $t \in \langle \mathcal{GS}_n \rangle$.

By induction we have $\mathcal{S}_n = \bigcup_{Z \in \mathcal{P}_n(Q)} \mathcal{E}(\pi_Z) \subseteq \langle \mathcal{GS}_n \rangle$. Therefore $\mathcal{S}_n = \langle \mathcal{GS}_n \rangle$. Since there are 2^{n-1} different $Z \in \mathcal{P}_n(Q)$, there are 2^{n-1} transformations in \mathcal{GS}_n . \square

Example 7. Suppose $n = 5$. The list of all transformations in \mathcal{GS}_5 is shown in Example 6. Consider $Z = \{3, 5\} \in \mathcal{P}_5(Q)$, and $t = [2, 4, 3, 5, 5] \in \mathcal{E}(\pi_Z)$. The transition graph of t is shown in Fig. 4 (a). As in Proposition 3, we have $Y = \{1, 2, 4, 5\}$, and $U = \{1, 2, 4\}$. To show that $t \in \langle \mathcal{GS}_5 \rangle$, we find $h_r = 1$.

Then $t' = \begin{pmatrix} 4 \\ 5 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = [1, 4, 3, 5, 5]$, and $Z' = \{1, 3, 5\}$. We assume that $t' \in \langle \mathcal{GS}_5 \rangle$; in fact, $t' = t_{Z'}$ in this example. We also need $Z'' = \{1, 1t\} = \{1, 2\}$, and $t_{Z''} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = [2, 5, 3, 4, 5]$. The transition graphs of t' and $t_{Z''}$ are shown in Fig. 4 (a) and (b), respectively. One can verify that $t = t't_{Z''}$, and hence $t \in \langle \mathcal{GS}_5 \rangle$. ■

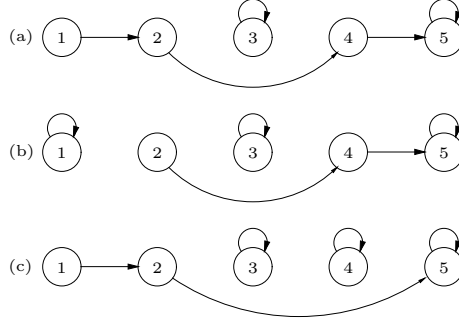


Fig. 4. Transition graphs of $t = [2, 4, 3, 5, 5]$, $t' = [1, 4, 3, 5, 5]$, and $t_{Z''} = [2, 5, 3, 4, 5]$.

Now, by Propositions 1, 2, and 3, we have

Theorem 4. *Let $L \subseteq \Sigma^*$ be a regular language with quotient complexity $n \geq 1$ and \mathcal{J} -trivial syntactic monoid. Then its syntactic complexity $\sigma(L)$ satisfies $\sigma(L) \leq g(n) = \lfloor e(n-1)! \rfloor$, and this bound is tight if $|\Sigma| \geq 2^{n-1}$.*

Remark 1. It was shown by Saito [16] that, if S is a \mathcal{J} -trivial submonoid of \mathcal{F}_Q , then $\Omega(S) = \{\Omega(t) \mid t \in S\} \subseteq \Pi_Q$ forms a \vee -semilattice, called a \mathcal{J} - \vee -semilattice, such that $\text{Max}(\Omega(t) \vee \Omega(s)) = \text{Fix}(t) \cap \text{Fix}(s)$. Let $\mathcal{P}_\vee(\Pi_Q)$ be the set of all \mathcal{J} - \vee -semilattices that are subsets of Π_Q . A maximal \mathcal{J} -trivial submonoid S of \mathcal{F}_Q corresponds to an maximal element P in $\mathcal{P}_\vee(\Pi_Q)$, with respect to set inclusion, such that $S = \bigcup_{\pi \in P} \mathcal{E}(\pi)$. $P \in \mathcal{P}_\vee(\Pi_Q)$ is called *full* if $\{\text{Max}(\pi) \mid \pi \in P\} = \mathcal{P}_n(Q)$, which is an maximal element in $\mathcal{P}_\vee(\Pi_Q)$ with respect to set inclusion. The monoid \mathcal{S}_n then corresponds to a full \mathcal{J} - \vee -semilattice, and hence it is maximal. Saito described all maximal \mathcal{J} -trivial submonoid of \mathcal{F}_Q and those corresponding to full \mathcal{J} - \vee -semilattices. However, here we consider the \mathcal{J} -trivial submonoid of \mathcal{F}_Q with maximum cardinality.

Remark 2. The number $\lfloor e(n-1)! \rfloor$ also appears in the paper of Brzozowski and Liu [6] as a lower bound and the conjectured upper bound for the syntactic complexity of definite languages. However, the semigroup B_n with this cardinality in [6] for definite languages is not isomorphic to \mathcal{S}_n , since B_n is not \mathcal{J} -trivial.

5 Quotient Complexity of the Reversal of \mathcal{R} - and \mathcal{J} -Trivial Regular Languages

In this section we consider *nondeterministic finite automata* (NFA's). An NFA \mathcal{N} is a quintuple $\mathcal{N} = (Q, \Sigma, \delta, I, F)$, where Q , Σ , and F are as in a DFA, $\delta : Q \times \Sigma \rightarrow 2^Q$ is the nondeterministic transition function, and I is the set of initial states. For any word $w \in \Sigma^*$, the *reverse* of w is defined inductively as follows: $w^R = \varepsilon$ if $w = \varepsilon$, and $w^R = u^R a$ if $w = au$ for some $a \in \Sigma$ and $u \in \Sigma^*$. The *reverse* of any language L is the language $L^R = \{w^R \mid w \in L\}$. For any finite automaton (DFA or NFA) \mathcal{M} , we let $\mathcal{M}^{\mathcal{R}}$ denote the NFA obtained by reversing all the transitions of \mathcal{M} and exchanging the roles of initial and final states, and by $\mathcal{M}^{\mathcal{D}}$, the DFA obtained by applying the subset construction to \mathcal{M} keeping only the reachable subsets. Then $L(\mathcal{M}^{\mathcal{R}}) = (L(\mathcal{M}))^{\mathcal{R}}$, and $L(\mathcal{M}^{\mathcal{D}}) = L(\mathcal{M})$. To simplify our proofs, we use an observation from [4] that, for any NFA \mathcal{N} without empty states, if the automaton $\mathcal{N}^{\mathcal{R}}$ is deterministic, then the DFA $\mathcal{N}^{\mathcal{D}}$ is minimal.

In 2004, Salomaa, Wood, and Yu [17] showed that if a regular language L has quotient complexity $n \geq 2$ and syntactic complexity n^n , then its reverse language L^R has quotient complexity 2^n , which is maximal for regular languages. As shown in [3] and [2], for certain regular languages with maximal syntactic complexity in their subclasses, the reverse languages have maximal quotient complexity. We now show that this also holds for \mathcal{R} - and \mathcal{J} -trivial regular languages.

First we consider \mathcal{R} -trivial languages. It was proved by Jirásková and Masopust [9] that, if L is an \mathcal{R} -trivial language with n quotients, then 2^{n-1} is a tight upper bound on the quotient complexity of L^R , and this bound can be met if L is a ternary language. Note that the syntactic semigroup of any \mathcal{R} -trivial language is a subset of \mathcal{F}_Q for some set Q . Hence the upper bound 2^{n-1} on $\kappa(L^R)$ can also be reached if L has n quotients with maximal syntactic complexity $n!$.

For \mathcal{J} -trivial languages L , it was conjectured by Masopust¹ that, if L has n quotients, then the upper bound 2^{n-1} on the quotient complexity of L^R can be reached using $n - 1$ letters. We now prove this conjecture.

Theorem 5. *For $n \geq 2$, if L is a regular \mathcal{J} -trivial language with quotient complexity $\kappa(L) = n$, then $\kappa(L^R) \leq 2^{n-1}$. Moreover, this bound can be met by a language L over an alphabet of size $n - 1$.*

Proof. Since any \mathcal{J} -trivial regular language is also \mathcal{R} -trivial, the upper bound 2^{n-1} also holds for \mathcal{J} -trivial regular languages.

To see that the bound is tight, consider the DFA $\mathcal{B}_n = (Q, \Sigma, \delta, 1, \{n\})$ such that $Q = \{1, \dots, n\}$ and $\Sigma = \{a_1, \dots, a_{n-1}\}$, where each a_i defines the following transformation of Q :

$$ja_i = j + 1 \text{ for } 1 \leq j \leq i - 1, ia_i = n, \text{ and } ja_i = j \text{ for } i + 1 \leq j \leq n.$$

DFA \mathcal{B}_n is minimal since, for each $i \in Q$, state i can be reached by a_{n-1}^{i-1} , and the word a_i is only accepted by state i . Let $L_n = L(\mathcal{B}_n)$. Then $\kappa(L_n) = n$.

¹ Personal communication

Let $\mathcal{N}_n = \mathcal{B}_n^{\mathbb{R}}$ be an NFA accepting L_n^R ; NFA \mathcal{N}_5 is shown in Fig. 5. Note that \mathcal{N}_n contains no empty states. Let P be any subset of Q containing n . If $P = \{n\}$, then it is the initial set of states of \mathcal{N}_n . Otherwise, suppose $P = \{p_1, \dots, p_k, n\}$, where $1 \leq p_1 < \dots < p_k < n$ and $1 \leq k \leq n-1$. Let $t = a_{p_1} \dots a_{p_k}$ be a transformation of Q . Then, for any $j \in Q$, $jt = n$ if and only if $j \in P$. Since $t \in \mathcal{T}_{\mathcal{B}_n}$, there exists a word $w \in \Sigma^*$ that performs the transformation t , i.e., $t_w = t$. This means that, for any $p \in Q$, $\delta(p, w) = n$ if and only if $p \in P$. Hence we can reach the set P of states of \mathcal{N}_n from the initial set of states by the word w . Since there are 2^{n-1} distinct subsets P of Q containing n , there are 2^{n-1} reachable states in $\mathcal{N}_n^{\mathbb{D}}$.

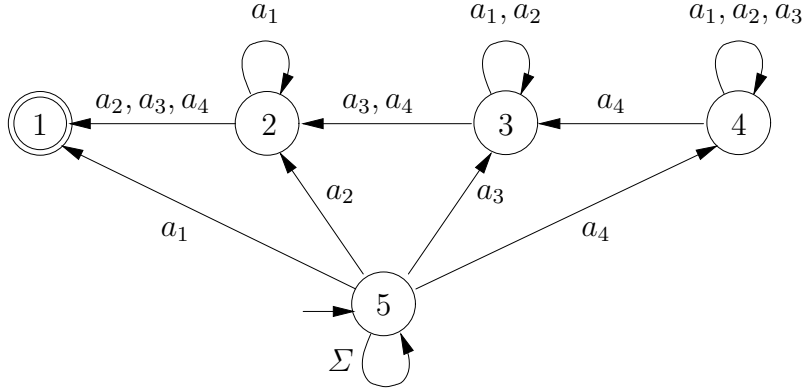


Fig. 5. NFA $\mathcal{N}_5 = \mathcal{B}_5^{\mathbb{R}}$ for $n = 5$ accepting L_5^R .

Since $\mathcal{N}_n^{\mathbb{R}} = \mathcal{B}_n$ is deterministic and \mathcal{N}_n has no empty states, DFA $\mathcal{N}_n^{\mathbb{D}}$ is minimal, and $\kappa(L_n^R) = 2^{n-1}$. This shows that the upper bound 2^{n-1} is tight for reversal of \mathcal{J} -trivial regular languages. \square

Consider again the above DFA \mathcal{B}_n . The orbit of each transformation a_i is $\{\{1, 2, \dots, i, n\}, \{i+1\}, \{i+2\}, \dots, \{n-1\}\}$; this is exactly the partition π_Z for $Z = \{i+1, i+2, \dots, n\}$. So $a_i \in \mathcal{S}_n$ by definition. Then the transition semigroup of \mathcal{B}_n is a subsemigroup of \mathcal{S}_n . It follows that, if a \mathcal{J} -trivial language L has n quotients and syntactic semigroup \mathcal{S}_n , then its reverse L^R has the maximal quotient complexity.

6 Conclusion

We proved that $n!$ and $\lfloor e(n-1)! \rfloor$ are the tight upper bounds on the syntactic complexities of \mathcal{R} - and \mathcal{J} -trivial languages with n quotients, respectively. When $n \geq 2$, the upper bound for \mathcal{R} -trivial languages can be met using at least $1 + C_2^n$

letters, and the upper bound for \mathcal{J} -trivial languages can be met using 2^{n-1} letters. It remains open whether the upper bound for \mathcal{J} -trivial languages can be met with fewer than 2^{n-1} letters. The syntactic complexity of \mathcal{L} -trivial languages is also open.

We also observed that, if \mathcal{R} - and \mathcal{J} -trivial languages have maximal syntactic complexities, their reverses have maximal quotient complexities. The proof of Theorem 5 can be extended to the following template for languages L in some subclass \mathcal{C} of regular languages: Suppose $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$ is the minimal DFA of L . To prove $\kappa(L^R) = f(n)$, where $f(n)$ is an upper bound on $\kappa(L'^R)$ for $L' \in \mathcal{C}$, one can show that there are at least $f(n)$ distinct subsets P of Q such that \mathcal{A} can perform a transformation t of Q with $it \in F$ if and only if $i \in P$.

References

1. Brzozowski, J., Fich, F.E.: Languages of \mathcal{R} -trivial monoids. *J. Comput. System Sci.* **20**(1) (1980) 32 – 49
2. Brzozowski, J., Li, B., Ye, Y.: Syntactic complexity of prefix-, suffix-, bifix-, and factor-free regular languages. *Theoret. Comput. Sci.* **449** (2012) 37 – 53
3. Brzozowski, J., Ye, Y.: Syntactic complexity of ideal and closed languages. In Mauri, G., Leporati, A., eds.: *DLT 2011*. Volume 6795 of *LNCS*, Springer Berlin / Heidelberg (2011) 117–128
4. Brzozowski, J.: Canonical regular expressions and minimal state graphs for definite events. In: *Mathematical theory of Automata*. Volume 12 of *MRI Symposia Series*. Polytechnic Press, Polytechnic Institute of Brooklyn, N.Y. (1962) 529–561
5. Brzozowski, J., Li, B.: Syntactic complexities of some classes of star-free languages. In Kutrib, M., Moreirad, N., Reis, R., eds.: *DCFS 2012*. Volume 7386 of *LNCS*, Springer (2012) 117–129
6. Brzozowski, J., Liu, D.: Syntactic complexity of finite/cofinite, definite, and reverse definite languages. <http://arxiv.org/abs/1103.2986> (March 2012)
7. Ganyushkin, O., Mazorchuk, V.: *Classical Finite Transformation Semigroups: An Introduction*. Springer (2009)
8. Holzer, M., König, B.: On deterministic finite automata and syntactic monoid size. *Theoret. Comput. Sci.* **327**(3) (2004) 319–347
9. Jirásková, G., Masopust, T.: On the state and computational complexity of the reverse of acyclic minimal dfas. In Moreira, N., Reis, R., eds.: *CIAA 2012*. Volume 7381 of *LNCS*, Springer (2012) 229–239
10. Klíma, O., Polák, L.: On biautomata. In Freund, R., Holzer, M., Mereghetti, C., Otto, F., Palano, B., eds.: *Third Workshop on Non-Classical Models for Automata and Applications - NCMA 2011*, Milan, Italy, July 18 – July 19, 2011. *Proceedings*. Volume 282, *Austrian Computer Society* (2011) 153–164
11. Krawetz, B., Lawrence, J., Shallit, J.: State complexity and the monoid of transformations of a finite set. <http://arxiv.org/abs/math/0306416v1> (2003)
12. Maslov, A.N.: Estimates of the number of states of finite automata. *Dokl. Akad. Nauk SSSR* **194** (1970) 1266–1268 (Russian) English translation: *Soviet Math. Dokl.* 11 (1970), 1373–1375.
13. McNaughton, R., Papert, S.A.: *Counter-Free Automata*. Volume 65 of *M.I.T. Research Monographs*. The MIT Press (1971)
14. Myhill, J.: Finite automata and the representation of events. *Wright Air Development Center Technical Report* **57–624** (1957)

15. Pin, J.E.: Syntactic semigroups. In Rozenberg, G., Salomaa, A., eds.: Handbook of Formal Languages, vol. 1: Word, Language, Grammar. Springer (1997) 679–746
16. Saito, T.: \mathcal{J} -trivial subsemigroups of finite full transformation semigroups. Semigroup Forum **57** (1998) 60–68
17. Salomaa, A., Wood, D., Yu, S.: On the state complexity of reversals of regular languages. Theoret. Comput. Sci. **320**(23) (2004) 315 – 329
18. Simon, I.: Hierarchies of Events With Dot-Depth One. PhD thesis, Dept. of Applied Analysis & Computer Science, University of Waterloo, Waterloo, Ont., Canada (1972)
19. Simon, I.: Piecewise testable events. In: Proceedings of the 2nd GI Conference on Automata Theory and Formal Languages, London, UK, Springer-Verlag (1975) 214–222
20. Yu, S.: Regular languages. In Rozenberg, G., Salomaa, A., eds.: Handbook of Formal Languages, vol. 1: Word, Language, Grammar. Springer (1997) 41–110